View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

Munich Personal RePEc Archive

Subjective Well-Being as Welfare Measure: Concepts and Methodology.

Fischer, Justina AV
OECD, Paris.

July 2009

Online at http://mpra.ub.uni-muenchen.de/16619/
MPRA Paper No. 16619, posted 29. March 2011 / 16:56

Subjective Well-Being as Welfare Measure:
Concepts and Methodology

Justina AV Fischer[1]
(ELS/SPD)

First version 6 January 2009,
revised version 27 July 2009

Abstract
Happiness research is on the rise, but is confounded by competing definitions of subjective well-being based on co-existing concepts, resulting in differing measures and giving rise to different potential policy applications.
This paper motivates the societal necessity for using well-being indicators and gives a short overview of the relation between the concepts 'subjective well-being', 'affect', 'life satisfaction', and 'happiness'. It describes their measurements and operationalizations in surveys, illustrates their philosophical roots, discusses their validity and reliability, and attempts to shed light on the scope of their policy applicability. Focus of this paper is on practical issues when applying measures of subjective well-being for policy evaluations. Target audiences of this paper are the interested public and laypersons, non-expert economists, and statisticians.

Keywords: Happiness, Life Satisfaction, Subjective Well-Being, Concepts, Methodology, Measurement, Econometrics, History, Behavioral Economics

JEL-codes: I31; D31; D01; D03; C81; C35; C13; B11

[1] (at time of writing), Dr. Justina AV Fischer, ELS/SPD, OECD, Paris, e-mail: javfischer@gmx.de.

Acknowledgement

Justina Fischer would like to thank Mark Pearson and seminar participants of the Max-Planck-Institute for Economics at Jena (Jan 2009) and the QEM-workshop of GESIS at Mannheim (15-16 July 2009) for helpful comments and suggestions. Particularly section 9 also draws from insights from the ISQOLS 2009 Satellite Meeting "Measuring subjective well-being: an opportunity for National Statistical Offices?" (July 23-24, Florence 2009) sponsored by the OECD, ISQOLS & ISTAT. This paper forms part of a project financed by the EU Commission.

The views expressed in this paper are those of the authors alone, not necessarily of the OECD or any of its Member countries.

Contents

1. Background: Why do we need a subjective well-being measure ?

1.1. Why income is so important to economists

Traditional economic theory equates utility with the satisfaction achieved through consumption of material goods, services, and leisure. In a life-time utility model, current income is split between current consumption and savings. Savings, however, constitute post-poned consumption, which will take place later in time.[2] Through mathematical transformation, utility can be thought of as being directly determined by income and the price level ('indirect utility'). As income increases, so does utility, so economic theory says. This view explains well the dominating role of national income and economic growth in public discussions and political decision-making processes.

The strong focus on consumption as sole determinant of satisfaction is much owed to the historical societal circumstances at the time when the foundations of economic theory were laid. In a Malthusian world which was characterized by low life expectancy, no education, high infant mortality, bad working conditions, but even higher fertility, most of the population was living at or below subsistence level (e.g. Courard-Hauri, 2007). In such world, there was an apparent connection between more income and higher well-being, easily measurable by the number of commodities and services one was able to consume (food, shelter, clothing, health care). Expressed in modern terms, the majority of the population lacked the 'basic needs' and the overall situation was not much different from what we observe today in developing countries. Assuming that utility could be directly observed and measured using survey questions on subjective well-being (SWB) (we return later to this point), and further assuming that GDP was an appropriate measure of population consumption, Easterlin (1974) has shown for the USA that there is no connection between the evolution of GDP and the development in average happiness.[3] (See also Fischer, 2009b, which presents a graphical presentation for a later and longer time horizon).

[2] Allowing for bequest, parents' savings may increase their children's' consumption and, thus, their children's satisfaction. Assuming interpersonal utility dependence (e.g. a dynastic utility) and perfect foresight, bequeathing their children will raise parental satisfaction.

[3] In the original article of 1974, Easterlin reports a decline in the share of happiest from 1963 to 1969, a time of unambiguously rising GDP per capita (15000 US$ - 18000US$), paralelled by an increase in the share of unhappiest, with a stable mean.

1.2. Social indicators needed to complement GDP

Seemingly, development of population well-being and evolution of wealth appear to be disconnected in developed countries (for a thorough discussion, see Fischer 2009b).[4] The level of national income from which on GDP and happiness are disentangled appears to lie between 10'000$ and 15'000$ per capita (e.g. Frey and Stutzer, 2000). In developed countries, GDP appears to lose its importance as single measure of societal progress because it neglects aspects of e.g.

- income inequality (social comparisons)
- social cohesion (commuting, mobility, migration),
- leisure time and other non-market activities (work-life balance)
- combating physical and social destruction, which enter positively (crime, terror)
- reduction of social capital, which enters positively (private market-priced services substitute voluntary engagement)
- deterioration of population health, which enters positively (health care costs)

In the so-called Istanbul declaration (2007)[5] various international and supranational organizations have acknowledged the need to complement purely consumption-based GDP with alternative measures of societal progress. Various indicators have been developed. The most prominent example is probably the Human Development Index, which is often applied for assessing the progress of societies in developing countries.[6] Other examples include, e.g. health indicators, equity indicators, prevalence of crime measures, etc. These indicators aim at reflecting the quality dimension of societal progress and are based on objectively measurable data. Notably, each of these social indicators relates to a specific dimension of human well-being. For example, of the social indicators presented e.g. in the OECD publication "Society at a Glance 2006" (OECD, 2007), the self-sufficiency and equity indicators relate to financial well-being, equity, the health indicators to mental and physical well-being, the social cohesion indicators to societal well-being, the sustainability indicators to environmental well-being, some social indicators to well-being in the family or work realm.

[4] One explanation is that with growing (national) income, through adaption processes, the consumption component of utility becomes less important, while social comparisons, which cancel each other out at the country level, gain in importance (Clark, Frijters, Shields, 2008; see also the overview in Fischer, 2009a).
[5] See http://www.oecd.org/dataoecd/14/46/38883774.pdf (9 Jan 2009)
[6] This index is a normalized indicator which is based on measures of life expectancy, literacy, educational attainment, and GDP per capita as its components.

1.3. A new social indicator: subjective well-being (SWB)

This report suggests and introduces a new social indicator, called 'subjective well-being' (SWB). 'Subjective well-being' is another expression for the perceived quality of the life one leads and the positive emotions one experiences; thus, in contrast to the previously discussed measures, it is based on purely subjective evaluations.[7]

The aim of such social indicators is to serve as benchmarks for policy evaluations, similarly to how GDP was (ab)used in the past.[8] In place of pondering whether policy A or policy B boosts economic growth or not, one may equally ask whether such a policy is conducive to people's health, reduces crime, raises social cohesion and increases networks, or, put simply, contributes positively to people's well-being.

The next section introduces the concept 'subjective well-being' in detail.

2. The concept of 'subjective well-being'

In general, definitions of subjective well-being are still in flux. They vary by researcher and by academic field. They equally vary by language in which scientific contributions are written. The following section makes an attempt to combine the differing, but overlapping concepts and definitions in a unified model, as it underlies most recent empirical happiness research. In general, the definitions developed by psychologists appear broader in scope compared to the concepts applied by sociologists and economists.

2.1. Psychological well-being

Psychologists term subjective well-being often 'psychological well-being' or 'positive mental health'. A typical definition of subjective well-being states that it is "about lives going well" which comprises a "combination of feeling good and functioning effectively" (Huppert, 2008). In the psychologists' view, the components of psychological well-being that relate to 'feeling good' include positive emotions (contentment, happiness), but also interest and engagement, confidence,

[7] As discussed later, the term 'subjective' is not derived from the fact that SWB is measured through self-report survey questions, but rather that its point of reference is internal (i.e. the definition of a 'good life' is up to the surveyed individual). In contrast, e.g. poverty is defined by an external reference point, such as a certain income level.

[8] Notably, statisticians always emphasize that the concept GDP started as a national account project and was never meant to assess societal progress.

and social relations. Those related to 'functioning' include having a goal in life, having control, developing one's potentials, and satisfying social relations. Similarly, the World Health Organization describes mental health/mental functioning as a concept which encompasses "subjective well-being, perceived self-efficacy, autonomy, competence, intergenerational dependence, and self-actualization of one's intellectual and emotional potential, among others" (WHO, 2001, page 5). This definition almost equals the quality-of-life definition (QoL), which recently appears to converge with concepts of psychological and subjective well-being (Camfield and Skevington, 2008).

2.2. Subjective well-being

Sociologists and economists define subjective well-being more narrowly, excluding some facets that are included by psychologists or the QoL approach. For example, physical and mental functioning would not be considered components of subjective well-being, but rather be viewed as their conditions. However, functioning (life-ability) would still be part of a broader concept of QoL. Excluded from the sociological and economic definition of subjective well-being are also aspects of meanings of life. The latter comprises both objective (outer qualities) and subjective (inner qualities) dimensions of 'life chances' and 'life results'. Within this concept, subjective well-being constitutes the subjective dimension of life results. The sociologist Veenhoven (2000) provides a detailed classification which is compared to competing popular concepts (and related measures) developed by various philosophers and social scientists (e.g. Brock, 1993; Sandoe, 1999; McCall, 1975; Ware, 1996; Ouelette-Kuntz, 1990; Cummins, 1993; Allardt, 1976), at the time of his writing. The following Table 1 illustrates these relations:

Table 1: Definition of subjective well-being (SWB)

	Outer Qualities	Inner qualities
Life chances	Liveability of environment (nature, society)	Life-ability of person (functioning, autonomy, control)
Life results	Utility of life (goal of life, moral perfection)	Appreciation of life (overall appraisals: subjective well-being, happiness, life satisfaction; domain satisfaction) * affective * cognitive

Source: Extracted from Veenhoven (2000), p.4 and p.11.

2.3 The affective and cognitive components of subjective well-being

'Subjective well-being' (Diener, 1984), the subjectively assessed quality of life from the beholder's viewpoint, is composed of two components: the affective and the cognitive component. The two components are considerably distinct: While the affective component constitutes momentaneous emotional states and instantaneous feelings, the cognitive component is defined as an ex-post, retrospective assessment of the quality of the life as a whole (Sumner, 1996). Thus, while the first component is related to experiencing the 'now and here', the second is 'remembered' and relates to the whole time horizon spanned by a human life up to the time point of evaluation. While the first constitutes an unfiltered physical reaction, the second requires the interviewee to step back and reflect, and, thus, is the outcome of a cognitive process. In general, the first component is often referred to as 'emotional well-being', 'affect' or 'hedonic flow of pleasures and pains', while the second is called 'life satisfaction' or 'happiness' (Andrews and Whitey, 1976). Confusion is often created by the fact that some researchers still use the term 'happiness' to refer to affective states (e.g. Diener et al., 2008). In the tradition of Kahneman (e.g. Kahneman and Krueger, 2006), the one, affective component would be termed 'experienced utility', and the other 'decision utility' or 'remembered utility'.[9] Table 2 summarizes these differences:

Despite their conceptual differences, both components of subjective well-being have in common not only that their points of reference are internal, which makes them subjective (in terms of comparison used), but also that they cannot be directly and objectively observed by an outsider. Most measures trying to grasp 'affect' or 'life satisfaction' are necessarily self-reported and self-evaluative, and, thus, subjective (in the sense of subjectively measured). In contrast, a health state or a personal financial situation can be objectively assessed, and an objective, external threshold (e.g. a poverty line) can be set. These two facets of subjective well-being are also reflected in the various variants SWB questions posed in surveys and utilized in empirical studies, as well as reflected in discussions on measurement and data collection issues. We return to this point of measurement in sections 6 and 7 on 'operationalization in surveys' and 'validity and reliability of SWB measures'.

[9] A similar concept to that of experienced utility is the notion 'process benefit' that derives from "direct subjective consequences from engaging in come activities" independent from the feelings about the outcome of this activity (Juster, Courant and Dow, 1985). Economists term these benefits 'procedural utility' as opposed to 'outcome utility'.

Table 2: Components of SWB: definitions, alternative terms, data collection

Affective component of SWB	**Cognitive component of SWB**
Definition and characteristics	
feelings and emotions pleasures and pains	the quality of life as a whole
instantaneous	ex-post, retrospective
hedonic	eudaimonic
event-specific	global
J. Bentham (Utilitarianism)	Aristoteles (eudaimonia <= eu + daimon)
Alternative labels	
Moment utility Hedonic utility Util Affect Affective Well-Being Experienced utility Procedural utility/ process benefits	Contentment Life satisfaction Happiness Cognitive Well-Being Remembered utility Decision utility/ outcome utility
Data collection	
Laboratory experiments, Day Reconstruction Method (DRM), Experience Sampling Method (ESM)	Opinion surveys (e.g. household panels)

Economic modelling of SWB

This distinction between affective and cognitive components of SWB is also reflected in some economic models. For example, Deaton (2007) assumes that (current) life-time utility is composed of an intertemporal integral of experienced utilities (utils) in the past. ("utility is intertemporally additive, which allows us to talk about instantaneous and lifetime utility in a simple way", p.7). Relating his model to the definition of subjective well-being, Deaton proposes that the cognitive component *life satisfaction* has the affective component as constituent. This view is in the tradition of J. Bentham (1789) who equally assumed subjective well-being to be the sum of pleasures and the absence of pain (hedonic experiences). However, Deaton (2007) is rather an exception, as the

focus of most economists lies on modelling life-time utility viewed as decision utility (as utility derived directly from consumption and saving over the life-time).

The next section discusses the relevance of the two components of SWB for policy-decision making from a SWB-conceptional viewpoint.

3. Relevance of affective or cognitive components for policy evaluations: adaptation and bounded rationality

There is an ongoing debate among happiness researchers whether empirical research should use, and policy recommendations should rely on, measures of the affective or the cognitive component of subjective well-being. In general, each component is useful in its own right, and appropriateness may depend on the policy question raised.

3.1. Affective well-being

Advantages

Laboratory experiments inflicting pain on test subjects have revealed the difference between experienced and remembered utility – the latter appears to be a weighted average of experienced utils. However, peak experiences and those at the end of the experiments appeared to be given larger weights compared to medium-impact or more distanced pains (so-called 'peak-end-rule'), while the duration of these experiences did not play a role for remembered utility (so-called 'duration neglect') (Kahneman et al., 1993). In other words, remembered utility is shaped by strong adaptation effects which occur almost instantly after a certain event has occurred. The longer the distance in time between event and assessment of personal SWB, the greater the difference between changes in affective and cognitive component of SWB will be. Fischer (2009a) discusses adaptation effects more in detail and provides examples for certain life events.

Policy applications

Thus, using measures of the cognitive component to assess the impact of e.g. pain and illness may actually understate their true effects on subjective well-being. However, the line of happiness research that assesses psychic costs of experiencing a certain event (such as physical damage, disability, unemployment) and attempts to calculate monetary (income) equivalents for their compensation may complement their life satisfaction studies with measures of experienced utility (affect) (e.g. Oswald and Powdthavee, 2006, for income equivalents for disabilities). However, assuming that subjective well-being is the intertemporal integral over utils, such approach requires the collection of affect measures on a more frequent basis and over a longer time span than most data bases provide (for a discussion of data collection, see below).

Applications of affect measures may also include an assessment of the impact of macro-shocks and their path of adaptation, particularly in case habituation occurs so fast that it may disguise (or have no) effects on remembered utility. For examples, the psychic costs of negative news such as the financial market crisis or terrorist attacks could be evaluated, and their adaptation paths could be observed. Is such news worse in their emotional response when they come as 'big bang', or when they are released step-by-step in homeopathic doses? The same question may be raised with respect to the optimal sequence of political and social policy reforms.

3.2. Cognitive well-being

The cognitive component determines choices

However, there is also a strong argument for focusing on remembered utility rather than experienced utility. Scientific evidence suggests that actual individual decisions and choices are determined by the cognitive component of SWB rather than its affective component, because of the above-described adaptation effects that take place (Kahneman and Krueger, 2006). To illustrate, our choice of future vacation is determined by how we remember the last vacation was, not what we experienced and felt at the time when we were in vacation the last time.

Basing choices on hedonism leads to sub-optimal behavior

Moreover, the overview article by Kahneman and Thaler (2006) presents arguments of why decision-making based on present-time hedonic experiences in place of remembered utility may

lead to choices at time point t = 0 which are not utility-maximizing in a long-term perspective, at time point t = 1. Assuming that consistency of preferences is a necessary prerequisite for utility maximization, any violation thereof (inter-temporal inconsistency, e.g. regret in t = 1 of choices made in t = 0) will prevent achieving maximal SWB in t = 1.

Prominent examples include the hungry shopper who shops excessively large quantities or unhealthy sweets, or the health club membership which is never utilized (so-called 'projection bias through anchoring in the current affective state').[10] Decision-making at time t = 0 for consumption in t = 1 may also be biased if the context of choice at t = 0 involves the comparison of two goods (joint evaluation), while later consuming only the chosen good. A typical example is the choice of an object in a store among a variety of similar, but not identical objects (e.g. television sets). The joint evaluation at t = 0 but separate consumption in t = 1 leads to a shift in weights attached to the good's characteristics for utility in t = 1. Similarly, simultaneous choice of several goods leads to a diversification bias, an over-evaluation of the importance of diversification (increased dissimilarity) for experienced and decision utility. Finally, hedonic adaptation in time t = 1 and thereafter is equally mispredicted, which plays an important role when deciding on major life events in time t = 0 (e.g. marriage, divorce, getting tenure, moving). In the long-run, only a few life events exert a lasting or permanent effect on utility as defined by the cognitive component of subjective well-being. However, mispredictions are less likely if t = 1 is close to t = 0, and if the individual can base her judgment on similar experiences made in the past.

Taken altogether, because the focus on experienced utility in t = 0 may bias predictions of future utility in t = 1, higher levels of actual utility in t = 1 could be achieved if choices were based on remembered utility instead. Possibly, in case personal remembered utility was not available to the individual for unbiased decision-making, an average person's utility changes may serve as benchmark. Hence, if policy-makers are interested in understanding people's actual choices and possibly, intend to influence their real-life behaviours to maximize their utilities, it is the cognitive component of subjective well-being that should be focused on for policy evaluations. [11]

[10] Even though Kahneman and Thaler (2006) claim that these biases relate to experienced utility only, they may actually affect remembered utility /life satisfaction. For example, the consumption of sweets in t = 1 may still create a positive affect, but a negative impact on remembered utility. So what is mispredicted in t = 0 (when sweets are bought) is not the affect in t = 1 (immanent utility experienced from consumption) but decision utility (remembered utility), e.g. when the gain in body weight becomes evident.

[11] Using the term 'decision utility' implies both analyzing the determinants of current subjective well-being as well as the impact of subjective well-being on future choices.

3.3. Affective versus cognitive well-being

However, depending on the type of well-being indicator employed, the effects for policy implications are unclear. The cognitive and the affective component may also relate to the short-term and long-term effects of circumstantial changes on SWB, which possibly can be influenced by policy-makers. In general, the variation in SWB across persons is said to be 30% determined by genes, thus not being subject to external interventions (Lykken and Tellegen, 1996). 30% of the variation is long-term, influenced by factors that can be changed only in the long-run (e.g. education, major life events), while 30% of the variation is short-run and momentaneous, caused by changes in mood and affect (Diener et al., 2008). Thus, policy makers may be able to influence up to 60% of the variance, but most likely deliberately only 30%. For example, long-term projects such as the improvements in institutional quality may be conducive to individuals' life satisfaction in the long run, but it may not affect the number of positive experiences or moods made on a daily basis. In the extreme case, better quality government institutions may even lead to less social interactions (due to less need to intervene from the citizenry's viewpoint) and thus to lesser occasions of experiencing positive feelings.

The following sections discuss the concept 'domain satisfaction' in relation with 'life satisfaction', before we turn to questions of operationalization of the two components of SWB using survey designs.

4. Life satisfaction and domain satisfaction

4.1. Definition and domains

Life satisfaction as a general assessment of the quality of one's life as a whole constitutes an overarching, multi-facetted, and multi-domain-encompassing concept. However, satisfaction can also be assessed with respect to certain aspects of one's life, or certain 'domains'. Prominent examples include financial satisfaction (= satisfaction with one's wealth and income situation), job satisfaction, satisfaction with one's health, or satisfaction with one's marriage and family life. Definitions of domain satisfaction vary by study and author. Cummins (1996) proposes the following seven domains of life as main contributors to quality of life: material well-being, health,

productivity, intimacy, safety, community, and emotional well-being.[12] However, besides these more traditional domains exist also governance structure and organisation (government, democracy, police), government services (welfare and community services), or life maintenance (housework, house maintenance).

4.2. Methodology and determinants of domain satisfaction

In principle, most of the methodological issues relating to life satisfaction, discussed below and above, are equally applicable to domain satisfaction measures. While life satisfaction is utilized and assumed to measure the latent concept 'utility', satisfaction with a specific domain of life is supposed to measure utility experienced in that domain only. In other words, while life satisfaction encompasses all market and non-market activities in the present and past without exception, domain satisfaction is derived from one outcome that relates to or occurs in that specific domain. Its determinants are, however, both domain-specific but also more generic. E.g. job satisfaction is determined by labor–related characteristics pay, work hours, work place characteristics, relations with colleagues, but equally by general socio-demographic factors that simultaneously relate to other domains of life such as gender, household type, education, and more.

Domain satisfaction and objective measures of domain-specific well-being

Subjective satisfaction with a certain domain is (potentially) mirrored by objective well-being measures relating to that domain; e.g. self-reported satisfaction with health may have its counterpart in indicators of objectively observable health status. Thus, various (but not all) domain satisfactions can also be related to objective social indicators that reflect the sufficiency of economic or financial resources, social well-positioning and integration in society, or a pleasant and good-health-supporting environment. In contrast, for the subjective measure of global life satisfaction, up-to-now, no objectively observable counterpart is available on a large scale. The section on the validity of subjective well-being measures (section 7) discusses in depth attempts to objectively measure people's utility.

[12] The same study uses information on more than 350 different domains.

Current research on/using domain satisfaction

The relation between domain satisfaction and life satisfaction appears to belong to the rather neglected aspects of happiness research. Indeed, most research either focuses on socio-economic determinants of satisfaction (with life or with a certain domain), or analyzes the impact of domain/life satisfaction on other economic outcomes (e.g. human behaviour such as early retirement decisions).

An example for the first type of study is Nielsen and Smyth (2008) which identifies determinants of job satisfaction of Chinese urban workers, such as gender, age, income, education, type of job, and migrant status. In a similar fashion, Ateca-Amestoy et al. (2008) explore the socio-demographic factors that contribute to satisfaction with leisure, among them also health and disability status, expenditures on leisure activities, social capital, and environmental factors. An example for the second type of analysis is the study by Rode et al. (2007), which predicts the intention to quit the job and finds that both job and life satisfaction independently serve as mediating factors of work-family-conflict.

4.3. Relation life satisfaction – domain satisfaction

Life satisfaction reflects an overall assessment of life in general, while domain satisfaction constitutes, analogously, an assessment of a certain life domain. Some happiness researchers believe that overall life satisfaction can be constructed using information on the various domain satisfactions. For example, using an objective measure of quality of life (the Comprehensive Quality of Life Scale, which the author also terms 'life satisfaction'), the meta-study by Cummins (1996) attempts to identify and quantify the contributions of seven domains of life to 'life satisfaction'. In contrast, other researchers believe that life satisfaction as a global evaluation cannot be broken into constituent parts.

That 68% of the variance in life satisfaction could be explained by eight domain satisfactions is reported by Fugl-Meyer, Branholm, and Fugl-Meyer (1991) for 200 inhabitants in Umea, a city in Northern Sweden. Factor analysis shows that emotion-related domain satisfactions (partnership relations, family life, sexual life, and absence of disability) account for about half of the variance, while satisfactions with contacts with friends (informal networks) and leisure activities account for

about 9%. The same contribution to life satisfaction is observed for the third factor, which is composed of vocational and financial satisfaction. The most well-researched correlation of a domain satisfaction with life satisfaction is the satisfaction with one's job, ranging between 0.3 and 0.4, depending on the study (see Rode et al., 2007, for an overview).

Table 3: Correlations of life satisfaction with domain satisfactions in the World Values Survey

	Categories	wave 5	waves 3-4	waves 1-4
Financial satisfaction	10	0.5379	0.6133	0.5800
Expected life satisfaction in 5 years	10			0.5744
Satisfaction with home life	10			0.5377
Feelings of happiness	4	0.4945	0.4795	0.4879
Job satisfaction	10			0.4278
Satisfaction with job security	10			0.3968
Do you ever feel very lonely	4			-0.3274
Ever felt depressed or unhappy	2			-0.2906
Self-reported health state	5	0.3196	0.2595	0.2955
Satisfaction with how democracy develops	4			0.2201
Satisfaction with people in national office	4		0.1020	0.1438

Notes: Correlation with life satisfaction which is measured on a 10-point scale. All correlation coefficients are significant at the 1 percent level.

4.4. Correlation of domain with life satisfactions for cross-country micro-data

The current 5th wave of the World Values Survey (2005), which contains data on 76'000 persons in 52 countries, allows to calculate the correlations between life satisfaction and several domain satisfactions, such as emotional well-being, satisfaction with the financial situation of the household, and subjective state of health. Also employed are the combined 3rd and 4th waves (1997-2001), comprising 120'0000 persons in more than 80 countries, and the combined 1st to 4th waves of the World Values Survey (1980-2001; 270'000 persons in more than 80 countries). These yield

additional insights into the unconditional correlations between various domain satisfactions and life satisfaction, reported in Table 3.

The general impression is that satisfaction with personal circumstances (job, family life, finances, positive affect and aspirations, $\rho = 0.4 - 0.5$) is stronger correlated with global life satisfaction compared to those domain satisfactions which relate to the politico-institutional environment (democracy, national office, $\rho < 0.2$). Inbetween are domains that relate to negative affective states and health (feelings of loneliness, depression, health state, $\rho = -0.3/\rho = 0.3$).

5. Excursus: philosophical roots of the notion 'happiness'

The present discussions on how subjective well-being is defined, how it can possibly be measured and, ultimately, what its determinants are, use arguments and concepts that have their roots in the Utilitarian philosophical tradition, which, in turn, is partly directly influenced by, and related to, the Classical Greek tradition. Utilitarian thought assumes that the ultimate goal of any human activity is to reach a state of individual happiness. Therefore, as J.S. Mill states, the rightness or wrongness of human actions can be judged by the proportion as they promote happiness – both with respect to individual happiness as well as with respect to general, societal happiness.

From Greek roots to Amartya Sen

That human beings strive for individual happiness has already been recognized by Plato and, more pronounced, Aristoteles. Plato viewed a happy life as one in which the three forces inherent in one's psyche were held in harmonic equilibrium. This equilibrium is maintained and supported by intellectual and practical activities in an ideal state, an equally in 'equilibrated' and 'psyche-equilibrating' governance structure (Plato, *Politeia*, second book), Aristoteles' concept was more real-world oriented. The latter defined a happy man as a man with 'eu-daimonia' (good daimon). This eudaimonic state could be achieved through fulfilment of human potentialities, by living out human reason through activities of human intellect and social-societal participation in a virtuous way. While Aristoteles' focus on the outliving of human capabilities constitutes a precursor to Amartya Sen's capabilities' approach of full human functioning (Sen, 1985; Nussbaum, 2000), his

focus on the quality of life as a whole as objective goal of one's life developed into one of the major strands of 18th century Utilitarism.

Utilitarism

In Utilitarism two concepts of 'happiness' co-exist: the hedonic one and the eudaimonic one. Rooting in Epicure's philosophy but re-introduced as concept by Jeremy Bentham (1789) into classical moral philosophy, the hedonic approach defines happiness as absence of pain and presence of pleasure (εδονη (gr.) = pleasure, joy). Any human action is deemed as good if, in a purely quantitative addition, the induced amount of pleasure is greater than the amount of pain. While the quantitative dimension takes account of both intensity and duration of pain, the quality dimension is neglected. J.S. Mill (1863), still in the Bentham tradition, makes a distinction between 'low pleasures' and 'high pleasures'. However, going beyond the hedonic conception, he assumes that happiness is more than a mere accumulation of positive emotional states and feelings, namely an ordered whole. Now in the eudaimonic tradition, he asserts that, ultimately, only higher pleasures such as pleasures of intellect, artistic activities, moral goodness, etc. can contribute to a happy life. Thus, Bentham is one of the few Utilitarians who tries to combine the eudaimonic with the hedonic tradition, of which the first relates happiness to life as a whole, while the latter focuses on instantaneous experiences of pleasure. This distinction between the eudaimonic and hedonic conceptions of a happy life still dominates the present-day discussions on subjective well-being (see Table 1), having an impact on, ultimately, policy conclusions.

We now turn to the question of measurement of SWB in surveys and appropriate data collection.

6. Operationalization of measuring subjective well-being in surveys

6.1. Measuring SWB components and data collection

Both cognitive and affective components of subjective well-being are measured through surveys in their various forms, including modern technologies of physical surveillance (e.g. measurement of electric resistance of skin to assess emotional states). The latter are rather cost-intensive when applied to a larger number of subjects or for a longer time span - therefore these are rather seldomly used, or exploited, by empirical happiness researchers. However, the cost argument also applies to

traditional paper-and-pencil-based surveys or telephone-administered questionnaires. At least in the past, this technological restriction lead to an annual cycle of repetition, if at all. Even though modern technologies may make surveillance of persons on a 24-hour basis affordable, this section will focus on what is available now for empirical analysis, which are people's answers to questions in more or less traditional surveys. We start with a general discussion of the relation between the subjective well-being component to be measured and suitable frequency of data collection, followed by a description of the most commonly employed survey methods and questions. Table 1 (above) gives a brief overview.

As described in the introduction to the concept 'subjective well-being', SWB encompasses two components: an affective component and a cognitive one. Important to the question of data collection is that the first relates, in principle, to the actual instant only (the 'now'), while the second implies a retrospection, starting from the 'now' backwards into the very past.

6.2. Data collection: affect measures

The affect component is related to instantaneous positive and negative feelings. A simple affect measure would be to ask "How do you feel (now/yesterday/last week) ?", "Do you feel lonely?" or "Do you feel depressed ?".[13] However, affect is usually more indirectly assessed by asking questions on recently experienced life events which are supposed to trigger negative or positive feelings, such as divorce, experiencing an accident, having been praised, or having successfully passed an exam (Bradburn, 1969). A widely employed affect measure is the so-called affect-balance-scale. [14]

It is obvious that the characteristics of instantaneity of affects, their 'now'-relatedness, makes any of their measures rather useless when applied too infrequently. What can we conclude when a person felt 'positive' at 23.45h on day 17 of October 2008, and then experienced a negative moment at 13.12h on day 20 in March 2009? Affects, according to their definition and very nature,

[13] According to most psychologists, asking about feelings experienced during the last week is a good proxy for the current affective state as *past* general feelings cannot be recalled so that the respondent projects her current affective state into the near past; on the other hand, asking about respondent's *present* general feelings leads to large measurement errors (personal communication, 24 July 2009).

[14] Bradburn's (1969) affect-balance scale is based on a calculation of the difference between positive and negative emotions. It is sometimes criticized for assuming that negative and positive effects oppose each other rather than being orthogonal (see also Fischer and Sousa-Poza, 2008).

change fast, possibly from one second to the next (psychologists claim that one's affective state changes, on average, every 10 minutes, personal communication, 24 July 2009). They constitute the most volatile and fluctuating component of subjective well-being, which are most sensitive to the tiniest environmental changes. They may even change due to entirely internal processes, such as sudden memories of past events (Schwarz and Clore, 1983).

Thus, longitudinal studies that employ affect measures should record them as frequently as possible. Traditional household surveys, however, are carried out on an annual basis. More appropriate are data collection methods requesting respondents to report their emotions and feelings several times during the day. The Experience Sampling Method (ESM) uses handheld computers to regularly self-assess subjects' affective states and record their current activities. The Day Reconstruction Method (DRM) allows at least for a daily recording by writing diaries. Usually, DRM asks participants to summarize their episodes (activities) of the preceding day (types and duration) and report the intensity of their feelings experienced during each of them (see also Kahneman and Krueger, 2006).

The Gallup data

It appears that the only institution that collects data on affective states on a daily basis for a large sample is the Gallup organization, which conducts such surveys on affective states of residents in the USA, with about 1000 representative individuals interviewed per telephone per day. In several European countries, a weekly collection will soon be realized (oral communication, 14 Nov 2008). Even though these data are not longitudinal but constitute repeated cross-sections, they will allow assessments of affective states for an average person, but within the limitations of micro-level cross-sectional design, explaining levels of affect. Alternatively, when aggregated to a higher geographical level (e.g. regional level), they may also allow to assess both levels and changes in affective states in the population.

6.3. Data collection: Life satisfaction measures

Life satisfaction measures constitute cognitive evaluation of the quality of one's life as a whole. By definition, life satisfaction should respond only to major life events, which affect the instantaneous utils over a long time-span in the same direction (assuming that lifetime utility is an accumulation of discounted instantaneous utils, e.g. Deaton 2007). Such events may include divorce and

separation, unemployment, becoming disabled, or losing a close person. Fischer (2009a) discusses the impact of such person-specific events on one's life satisfaction. In theory, for the general case that no life event occurs between two points in time, t = 0 and t = 1, it should not matter whether life satisfaction is measured in t = 0 (now) or in t = 1 (e.g. 20 minutes later, today, or one day later). (Notably, the *theoretical concept* 'life satisfaction' is distinct from the *survey question* on life satisfaction. In real life, we observe that affective states do influence how the life satisfaction *question* is responded to, giving rise to measurement errors in the responses. We discuss this point in the section on test-retest reliability in section 7.)

Up-to-date, most empirical studies still rely on survey questions which are posed (if carried out as longitudinal survey) in regular intervals, usually annually – an appropriate frequency for assessing (changes in) the 'life satisfaction' component of SWB.[15] Most national household panel data are collected annually, as are some international surveys that constitute repeated cross-sections (e.g. International Social Survey Programme, the European Barometer Survey).[16] (Technical note: While in a household panel a representative, but identical sample of the resident population is interviewed on a regular basis, a repeated cross-section is characterized by the fact that the composition of the sample changes from one wave to the next, although each wave remains representative for the underlying national population.) As exception, the Australian household panel appears to be conducted 3-monthly which allows more precise evaluation of adaptation effects on life satisfaction over a longer time span.

6.4. Happiness with life versus satisfaction with life

The cognitive component of SWB, 'life satisfaction', is assessed using two types of survey question: one which asks the respondent to rate the 'happiness' with her life (or some variant), while the alternative and recently more commonly used survey question requires an assessment of the respondent's overall 'satisfaction' with her life as a whole (or some variant). Let us term the first type the 'happiness question' and the second the 'life satisfaction question'. Notably, both aim at measuring the latent *concept* life satisfaction.

[15] Annual surveys of life satisfaction are contained in the British, German, and Swiss household panels. The Australian household panel is updated on a quarterly basis. See Fischer (2009a) for the discussion of adaptation effects.
[16] The famous World Values Survey (WVS) is collected only irregularly, on average every 5 years, with an altering number of participating countries.

Examples for proto-type survey questions are “Overall, how satisfied are you with your life these days?”, or “How happy are you with your life?”. Traditionally, the answers to the happiness question express the degree of agreement, e.g. ‘strongly disagree’, ‘somewhat disagree’, ‘somewhat agree’, etc.), and is measured on a 3- to 4-point, or sometimes 5-point scale. Instead, the life satisfaction question is recorded as 10-point or 11-category variable. Possible answers are presented to the interviewee on a numerical scale, with the lowest number labelled ‘completely satisfied’ and the highest ‘completely dissatisfied’ (or some variant thereof) and no labelling for values inbetween. (In general, items asking the respondent to express the degree of agreement with a certain statement, as the happiness and some affect measures do, are said to use the so-called *Likert-scale* (Likert, 1932); psychologists claim that asking for respondent’s ‘agreement’ triggers a ‘yes’-bias, personal communication 15-16 July 2009). Please note that the life satisfaction question/item is often said to be measured on the satisfaction-with-life-scale (SWLS), in terminal analogy to the affect-balance-scale that measures affects (Diener et al., 1985)). How the differences in scaling influences the choice of method of analysis and interpretation of outcome is discussed in the section 8 on ‘Cardinal versus ordinal interpretation of SWB measures’.

The Cantril’s ladder scale of life satisfaction

Recently, the Cantril’s ladder version of the life satisfaction question has become more and more popular, most prominently applied by the Gallup organization. The Cantril-ladder version (first described in Kilpatrick and Cantril, 1960; Cantril, 1965) requires respondents to rate the satisfaction with their lives relative to self-defined minima and maxima of life satisfaction (concepts of “the worst/best possible life for you”, applied to the question “on which step of the ladder do you feel you personally stand at the present time”). This question may well abstract from affects through self-anchoring, on the one hand, and well solve the ‘closed-scale problem’ of the traditionally employed life satisfaction and happiness questions described above.[17] However, the analysis by Bjørnskov (2008) suggests that the World Values Survey question and the Cantril’s ladder question measure two distinct concepts of life satisfaction.[18] In particular, the concept of a ‘best possible life’ may be sensitive to culture and expectations, for example on the economic development. Indeed, according to his analysis, correlations of life satisfaction à la Cantril with economic development are quite low, and the famous positive 'Latin America' effect disappears.

[17] The ‘closed-scale problem’ refers to the fact that levels of life satisfaction beyond the maximum of the scale (e.g. higher than category 10) cannot be recorded.

[18] Using the Cantril’s ladder in the Gallup data also implies a downsizing of country averages compared to the equally scaled World Values Survey life satisfaction question, in general.

Even worse, in a multivariate analysis democratic rights, corruption, and political ideology show reversed signs compared to the well-known findings based on the traditional happiness and life satisfaction questions. However, more research is needed that clarifies the differences between the various variants of the SWB question.

Separation of measures relating to two distinct components of SWB

From a conceptual perspective, since the affective component of SWB is assumed distinct from the cognitive one, measures relating to the first should not overlap with measures of the second. In other words, one may well feel 'unhappy' but still overall assess the quality of one's life as 'good'. To illustrate, the fact that I am drinking a cup of coffee now with my friends (which creates a positive affect) should not have an influence on how satisfied I am with my life as a whole. Indeed, Helliwell (2008) shows that average affect is fairly identical across regions, in contrast to average life satisfaction which varies substantially in the same sample. Shimmack, Schupp and Wagner (2009) show that the correlations of personality traits or economic factors with SWB differ substantially depending on which component is analyzed.[19]

Similarly, due to the retroperspective and evaluative nature of both the happiness and the life satisfaction questions, in theory, respondent's answers should not be impacted by changes in her current mood. As discussed later in section 7 on reliability, this is, however, not the case: answers to both questions are well impacted by affective state, a typical case of measurement error. Indeed, the correlations between affective state measures and measures relating to the cognitive component of subjective well-being even serve as validity test in some studies.

At this point of discussing advantages and disadvantages of several variants of the subjective well-being question, attention should be drawn to the fact that the degree of contamination by affective state does most likely depend on differences in wording. In particular, it can be suspected that in some languages the word 'happy' in the 'happiness question' variant of the decision utility measure is likely to make answers more responsive to moods than the 'life satisfaction'-variant (e.g. Clark and Lelkes, 2009). However, this statement still awaits its empirical prove.

[19] Correlations of personality traits are larger with affect while those of socio-demographics are larger with life satisfaction. Regressions for decision utility suggest that personality traits only approximate (omitted) affect.

7. Validity and reliability: prerequisites for interpersonal and cross-country comparisons

Both validity and reliability determine the quality of statistical measures. Simply said, reliability relates to the fact that a measure yields consistent results, while validity means that the chosen measure actually assesses what it is supposed to assess. Validity is assessed by conducting empirical tests that examine whether the measure behaves as would be expected according to the theories that underlie the construct (e.g. whether respondents to the life satisfaction question report as predicted by microeconomic utility theory).

Reliability and validity are interlinked: the first is a component of the second. Both validity and reliability are necessary to establish interpersonal comparability of measures of SWB, and thus for exploitation of survey-based well-being questions for policy assessment in OECD countries.

7.1. Reliability

Reliability of a measure relates to issues of consistency, which is usually hampered by measurement errors. Measurement errors occur when e.g. in a questionnaire unintentionally the wrong box is checked, or the question is misunderstood by the respondent. Measurement errors may also occur when the ordering of the questions in a survey influences the answers to a certain question because, e.g. memories or associations have been triggered by preceding questions.[20]

How to test reliability

Reliability is currently examined using three methods: first, by posing the same question twice during the same interview. Second, by comparing answers to two variants of SWB measures, for example the life satisfaction and the happiness questions, or by comparing a single-item question with a multiple-item-scale question (Schwarz and Strack, 1999).[21] Finally, test-retest correlations are calculated over a longer time span, which eliminates memory effects but makes actual difference in well-being due to actual circumstantial changes more likely.[22]

[20] Based on this experience, social scientists suggest that SWB should be assessed in surveys right at the beginning, prior to asking about respondent's personal circumstances (e.g. income, marital status, etc.), and the cognitive component prior to the affective component (personal communication, 15-16 July 2009).

[21] This test may underestimates reliability if people think they are expected to answer differently to two seemingly different questions.

[22] This test may overestimate reliability due to recall bias, namely that persons remember how their answers to the question when it was posed for the first time, and wish to be consistent.

Results

The first and second reliability tests yield correlations between 0.6 and 0.7 (e.g. Wanous and Hudy, 2001; Schimmack and Oishi, 2005; Lucas, Diener and Suh, 1996; Kahneman et al., 1993).[23] In panel studies, reliability estimates are higher (up to 0.7, Ehrhardt et al., 2000) compared to cross-sectional studies (six-week re-test correlation of around 0.55, Schimmack et al., 2009), arguably due to interviewees' increased experience with surveys (e.g. Schimmack et al., 2009).[24]

Reliability may vary by a) the variant of SWB measure employed (more cognitive or more affective) and b) the scaling that is used. In general, reliability appears less sensitive to whether life satisfaction, happiness or affect was measured rather than to the number of categories. 10- and 11-point scales were found to produce more reliable answers (by up to 20 %) compared to 7-, 5- or 4-point scales (e.g. Kroh, 2009; Saris et al., 1998).[25] However, compared to objective measures of economic outcomes, such as income (rho = 0.9), test-retest correlations are considerably lower.

7.2. Validity

Causes for loss of validity

Validity implies that the underlying construct (unobservable subjective well-being) maps well into the measurement space (e.g. the life satisfaction question). There are several causes that could distort the relation between SWB and any of its survey-based measures. Notably, some of these concerns can be addressed through the choice of type of data and statistical methods, others can't.

- As Diener et al. (2008) argue, first, if answers were driven by external factors only they might have no relation at all with the to-be-measured internal state. For example, cultural norms and 'moral visions' may constrain the range of 'feasible' answers so that they would reflect true subjective well-being only within a certain range. Examples are the 'modesty/self-enhancement biases' in collectivist/individualist cultures (e.g. Christopher, 1999).

[23] The time gap between the test and the re-test ranges between 2 weeks and 6 weeks in these studies.

[24] Based on this experience, some social scientists suggest that the first waves of a household panel should not be included in any empirical analysis of SWB because of the still-ongoing learning process of the interviewees. In particular, interviewees are said to at first report SWB levels too high which they adjust downward in preceding waves (personal communication, 15-16 July 2009).

[25] Kroh (2009) also shows that reliability is not affected by whether data are collected in form of interviews or as paper-and-pencil questionnaires.

- The chosen measure may reflect a subjective state but not the one it is supposed to measure. For example, the life satisfaction question aims at assessing one's general evaluation of the quality of life as a whole. Given this, influences of emotions triggered by e.g. the weather or other immediate context undermine the validity of the life satisfaction question in that respect (Schwarz and Clore, 1983).
- In this context, personality traits (e.g. negative or positive affectivity, neuroticism, extraversion) may equally restrict the validity of this measure; the role of personality traits for SWB assessment has been demonstrated in studies on mono- and diacygotic twins (Stubbe et al., 2005; Lykken and Tellegen, 1996); personality traits influence the probability of recalling positive life events or to report signs of depression (Layard, 2005).[26]
- It is often argued that words or numbers, even more on a closed scale, are inappropriate tools to express internal states of subjective well-being. A similar problem relates to issues of insufficient translation of labels from one language to another, although Ouweneel and Veenhoven (1991) claim that differences in language do not cause incomparability.
- Depending on what one aims at measuring, also transitory influences on SWB through certain life events (see Fischer, 2009a) could be regarded as threat to the validity.

Face and content validity

In general, so-called *face* and *content validity* are of no concern for commonly employed SWB measures: as e.g. the life satisfaction question asks for a global assessment of one's quality of life, it encompasses all aspects and domains of the underlying construct 'subjective well-being'. In contrast, affect measures that are based on several items may not reflect 'happiness' in its full breadth (e.g. the so-called Bradburn's (1969) affect-balance-scale had to be tested in that respect). In this sense, measures of e.g. self-esteem and optimism do not constitute valid measures of either global life satisfaction or affect because they simply reflect personality traits ('internal dispositions' as called by psychologists), but not experiences (experienced utility).[27]

[26] According to Schimmack et al. (2008) particularly neuroticism and extraversion are main predictors of affect as these traits relate to how affective information is processed in the human brain.

[27] Using this related, but distinct concept of personality trait so-called *discriminant validity* of SWB measures can be assessed; discriminant validity assesses to what extent two related concepts (personality and SWB) can be measured as distinct concepts (e.g. Lucas et al., 1996)

.

Convergent validity

More important for assessing SWB measures may be the so-called *convergent validity*. Most of the published validity studies appear to relate to this facet of validity. *Convergent validity* assumes that if two measures assess the same underlying construct, they must be considerably correlated. If one of these measures has been proven valid, the alternative measure's validity can be evaluated based on its correlation with the first.

- Although they focus on different components of SWB, questions on affect and global life satisfaction are both supposed to measure the latent construct SWB. They appear considerably correlated, not only across individuals, but even more when country means are calculated (Schimmack et al., 2002).
- Convergent validity can also be tested evaluating respondent's SWB by external informants who provide an independent, but subjective assessment of the interviewee's SWB, knowing her personal circumstances and preferences (e.g. Wilson, 1967). Correlations are about 0.4, and convergent validity seems to be higher for the life satisfaction question than for affect measures (Schneider and Schimmack, 2009). Some researchers view the congruence between observed facial impressions of 'happy' persons and their self-report affect/life satisfaction as form of such external validation (for an overview, see Kahneman and Krueger, 2006).
- Validity tests which are based on objective, since physical, measures include those reporting large positive correlations with stress resistance, lower probability of getting a cold, and shorter recovery time for wound healing (Layard, 2005; Cohen et al., 1999; Kiecolt-Glaser et al., 2002). Similarly, assessing psychophysiological processes in the brain may help overcome the biases through labelling and translation. Up to now, the few neurological studies that exist support only a weak correlation (e.g. Van Reekum et al., 2007; Urry et al., 2004).[28]
- External clinical validation studies using objective measures of health have been carried out by Blanchflower and Oswald (2008a) for high blood pressure (hypertension) and cardiac diseases. They find a considerably high correlation of country rankings using either the

[28] Brain activity tests relate to the affective component of well-being. Emotional well-being was found to correlate with activity in the left prefrontal cortex, and unpleasant experiences with the prefrontal cortex on the right. Correlation of SWB question with the differences in both sides activities was 0.3 (Urry et al., 2004).

objective or the subjective measure of SWB. Similar validity studies exist for heart rate among men and their cortisol levels and cardiovascular behavior (Steptoe et al., 2005).[29]

- An alternative objective measure of life satisfaction may be the occurrence of suicide by the very unhappy, which we discuss in a separate section.

7.3. Happiness and suicide

According to the economic model by Hamermesh and Soss (1974) a rational life utility-maximizing individual puts an end to her life if the discounted (expected) net-utility reaches zero (or falls below). Based on this assumption, persons who are at risk should be found at the lower tail of the life satisfaction distribution (current net utility > 0 but small), and an even minor negative life event may balance out costs and benefits of continuing life. Both Helliwell (2007) and Daly, Wilson, and Johnson (2007) find that socio-economic and institutional factors exert influences in suicide models comparable to the effects observed in happiness models. However, these studies establish validity only if suicide committers are assumed to be rational, and one may still argue that those persons behave systematically differently compared to the rest of the population (mental illness).

However, Koivumaa-Honkanen et al. (2003) show in a study which followed 30'000 individuals over a time span of 20 years, that baseline happiness is a good predictor of committing suicide later in life, with suicide risk increasing by roughly 2 percentage points when moving down from any happiness category to the next lower (see also Weitoft and Rosen, 2005, for affective well-being). One of the most convincing validity studies is that by Daly and Wilson (2008) – both methodologically and sample size-wise. They compare the individual risk of falling below a certain threshold between life satisfaction and suicide models, each based on micro-level data from the US, from 1970 to 2006, where drop-outs due to suicide can be identified. They find that income-related and other socio-demographic life events shift people's position on the happiness continuum identically, whether measured in terms of objectively measured suicide risk or subjective risk of ill-being (= being in the lowest category of the happiness scale).

[29] Validity of country rankings based on subjective health state measures is equally confirmed (Blanchflower and Oswald, 2008b; Mojon-Azzi and Sousa-Poza, 2007). Blanchflower and Oswald (2008b) give references from the medical literature that suggests that self-report hypertension is a valid measure of actual disease.

Table 4: Convergent validity tests

	Source (examples)
Correlation of different SWB questions Test-retest correlations	Schwarz and Strack (1999) Wanous and Hudy (2001) Schimmack et al. (2009)
External informant's rating	Wilson (1967) Schneider and Schimmack (2009) Kahneman and Krueger (2006, overview) Layard (2005, overview)
Neurological tests of brain activity	Van Reekum et al. (2007) Urry et al. (2004)
Objective measures of health (heart rate, blood pressure)	Blanchflower and Oswald (2008a) Steptoe et al. (2005)
Objective measures of health improvement (recovery from cold, wound healing)	Cohen et al. (1999) Kiecolt-Glaser et al. (2002)
Suicide research (macro-level, micro-level)	Helliwell (2007) Daly and Wilson (2008)

7.4. Validity of aggregate measures of life satisfaction

Validity is also established for aggregate measures of subjective well-being, such as national means of happiness scores, or the shares of those reported in the highest or lowest categories. The afore-mentioned study by Blanchflower and Oswald (2008a) compares the inverse country ranking based on objectively measured blood pressure with that obtained from aggregate subjective well-being and finds a substantial correlation. The Spearman rank correlations range between 0.5 and 0.6.[30] OECD (2007, p.106-107) reports a negative correlation between suicide rates and mean life satisfaction. Helliwell (2007) finds that aggregate factors, such as formal and informal institutions, but also macro-economic condition, influence the occurrence of suicide and population well-being in qualitatively identical ways. Similarly, Daly and Wilson (2008) report simple correlations of socio-economic population characteristics with suicide rates and the share of unhappiest in the US population that are qualitatively comparable.

[30] More specifically, country rankings are based on country dummy coefficient estimates both in a hypertension model and a life satisfaction model (4-point scale), using the European Barometer data of 2001 on 15'000 persons in 15 European countries.

Table 4 (above) briefly summarizes the methods used to assess the convergent validity and reliability of survey-based subjective well-being measures.

7.5. Construct validation: predictions versus empirics

Construct validation implies the assessment of the quality of a specific measure based on whether predictions derived from (economic) theory on the underlying construct can be verified when this measure is employed in empirical analyses. For example, if the SWB measure claims to approximate decision utility (as underlying construct), and if economic theory predicts that utility rises in consumption, then any empirical study of the relation of individual consumption with this SWB measure should yield a positive correlation.[31] A positive correlation (even with decreasing marginal returns) would then be interpreted as indicator that the measure was a valid representation of the underlying construct. Fischer (2009a) discusses in depth to what extent micro-economic predictions can be replicated in empirical happiness research.

Empirical construct validation has been achieved for the predicted effects of individual income, but also health state, unemployment, divorce, disability (as proxies for income loss), education (as proxy for income)[32] (see Dolan et al., 2008, for a recent literature review from an economist's perspective).

In principle, any misguided theory on what determines utility may yield wrong predictions, for which no corroboration with real-life data can be found. For example, construct validation has failed what gave rise to the so-called Easterlin (1974) paradox. Similarly, the U-relation between SWB and age is considered as well-established fact, but economists have failed in modelling this relation using economic theory (for discussion, see Fischer, 2009c). In fact, many failures to empirically corroborate theoretical predictions may cast doubt rather on the variant of economic theory used rather than on the validity of the measure.

The role of research design

[31] According to economic theory, direct utility, which depends on consumption (with income as budget constraint), can be re-formulated as so-called indirect utility that is a function of income (and price vector) only. Indirect utility rises in income.

[32] Notably, these associations do not exclude other, non-economic explanations for their effects. However, in the light of construct validation which tests the consistence of empirical results with one specific theory, alternative explanations do not matter. See Fischer (2009a) for additional non-economic explanations for some of these effects.

Some of the concerns about the reliability and validity of subjective well-being measures can be addressed by choice of method of analysis. Measurement errors that may affect the reliability of SWB measures can by addressed by using large samples, including several thousands of respondents. Using large samples enhances the precision of statistical analyses. Psychologists often address measurement errors in smaller samples by basing the item 'SWB' on several survey questions and using statistical methods to construct/reveal the underlying latent variable (e.g. principal component analysis) (Schimmack and Oishi, 2005). Some surveys randomize the ordering of the questions across interviewees in order to prevent the bias through order effects.

7.6. Fixed effects account for personality traits

Other concerns about the validity of subjective well-being measures can also be addressed by choice of method of analysis. Notably, each 'cure' comes along with other, new disadvantages that have to be taken into account.

The confounding influence of personality traits, giving rise to potentially spurious correlations between two subjective measures, such as self-assessed health and SWB, can be addressed through the use of household panels (longitudinal studies) that observe the same person over some years. These panels allow to directly partial out the effects of personality traits so that partial correlations between SWB and its determinants (correlations conditional on these unobserved individual characteristics) can be obtained. Technically speaking, this can be achieved through calculating first differences (relating changes in SWB to changes in its determinants) or inclusion of so-called individual fixed effects. Indeed, Ferrer–i-Carbonell and Frijters (2004) have shown that the bias by not taking into account unobserved personality traits is more severe by far compared to the bias that is generated through the violation of the cardinality assumption.[33] However, these methodological approaches have the disadvantage that they partial out the effects of *all* time-invariant factors likewise, not only those of personality traits in specific, so that their influence can rarely be tested. Typical examples are gender, religious affiliation, mother tongue and childhood experiences. Education at school is completed by most people before they enter the panel (not prior to the age of 16). Thus, for most surveyed persons, education is part of their time-invariant characteristics. Other examples pertain to occupational and marital status – statistical identification of their impact on SWB rests on those individuals who change their status during the period of observation.

[33] In principle, estimation with OLS when the dependent variable is ordinal (measured in categories) violates the cardinality assumption of OLS.

Fixed effects account for culture bias

The inclusion of fixed effects may also partly resolve the issue of culture and language bias, which may affect responding behaviour. Language, ethnicity, and culture biases may occur at the macro-level, namely as trait of the country one lives in, as well at the micro-level, as individual trait. Well-known examples are the collective optimism bias in South-American countries and the national culture of misery and critical self-reflection in France (e.g. Duncan, 2005).[34] In a household panel, individual fixed effects take account of both types of cultural biases (e.g. Ferrer-i-Carbonell, 2005, for an example), while excluding them calls for direct control for ethnicity (for the cultural divide in Switzerland, see Dorn et al., 2008). However, most cross-national comparisons still rely on cross-sections of micro-data (or repeated cross-sections thereof (e.g. Blanchflower and Oswald, 2008a, b, for examples). In such setting, country fixed effects can de-bias the estimates for country-specific culture effects, but not for those that may occur across individuals (of the same country). These must be resolved by inclusion of direct controls (if available) for personality traits, mother tongue or ethnicity, and religion.

Validity of aggregate measures of happiness

Country rankings based on aggregate measures of subjective well-being appear more robust to differences in national cultures. Particularly the validity tests by Blanchflower and Oswald (2008a, b) suggest that the happiness rankings can be fairly reproduced when measures of respondents' health in place of SWB measures are employed. However, they cover only 16 European countries. Convergent validity is also considerably stronger for aggregate measures of the cognitive component of SWB in terms of the correlation between the 'happiness' and 'life satisfaction' questions, assuming that both intend to measure the same underlying construct. At the country level, correlation is about 0.7 for more than 80 countries of the 3rd and 4th waves of the World Values Survey. In contrast, at the individual-level, correlation in the pooled sample amounts to only 0.4, ranging between 0.0 and 0.8 in the country samples.

However, even though the strong validity of aggregate measures may be appealing to researchers and policy-makers, using them comes at a cost. For analysis, having an aggregate outcome as explanand requires aggregation of the explanatory variables. However, through the aggregation process important individual-specific information is lost, which affects the outcome of the analysis

[34] Similarly, East-European residents report systematically lower happiness, which is, most possibly, an effect of institutional and economic transition and therefore not a time-invariant cultural trait.

that utilizes such measures. Not only does aggregation make it impossible to disentangle certain micro-level from macro-level effects, neglecting their interactions (e.g. for unemployment), but even worse, the loss of information may bias the results at the aggregate level compared to the micro-level (Robinson, 1950). Fischer (2009a, b) discusses this issue in detail and gives specific examples for OECD countries.

7.7. Future development for cross-country comparability of SWB measures

Recently, efforts have been undertaken to make individual responses to variants of the happiness question comparable across countries and cultures. There are two approaches: the first attempts to determine 'conversion' rates between two respondents with differing cultural backgrounds. The second attempts to equivalize the definitions and concepts of a 'good life' across persons that are implicitly assumed in the happiness question.

Happiness Scale Interval Study

Using concepts from sociology and psychology to measure latent constructs, respondents are surveyed in countries around the world. In the Happiness Scale Interval Study led by Prof. Ruut Veenhoven (University of Rotterdam, Veenhoven, 2009) persons from different cultures and countries assess the range of the happiness categories on the (less-language-sensitive) life satisfaction scale, so that the ‚culture' bias can be determined and a country-specific conversion rule can be calculated (personal communication with Ruut Veenhoven, January 2009).

Vignettes/ factorial survey design

The cultural bias is also mitigated by using a so-called vignette design. Such design aims at replacing abstract culture-specific values (e.g. 'best life') with a description of concrete events and situations, which are, supposedly, less susceptible to culture-specific perceptions (e.g. of what a good and desirable life constitutes). Because the respondent is put in a specific, but hypothetical situation, such question design is also referred to as 'quasi-experimental'; it is also referred to as 'factorial survey design' since abstract values are replaced by a factual situation. Given that the chosen vignettes are reasonably simple, translation incompatibilities should only play a minor role. Most recent cross-national surveys such as the World Values Survey, the International Social Survey Programme, the European Values Survey, or the European Barometer Survey do not employ

vignettes. The Gallup organisation plans to apply vignettes to the Cantril-ladder question of life satisfaction.

8. Cardinal versus ordinal interpretation of SWB measures

8.1. Types of survey questions

Survey questions on SWB (e.g. happiness or life satisfaction question) measure the respondent's answer on a scale of integers. The happiness question usually employs 3 to 5 categories, while life satisfaction is usually recorded on a 10- or 11-point scale. Both are, in principle, ordinal rankings of SWB (e.g. worst, less than worst, good, best), calling, in principle, for statistical methods that take the ordinality of the scale into account. Ordinal rankings imply that moves from one category to the next constitute only qualitative statements (improvement/worsening, respectively) and that changes by one category are not quantitatively comparable across their starting levels. In contrast, cardinality assumes that moving from e.g. 3 to 4 is an improvement as large as that when moving from 4 to 5, and that moving from 2 to 4 is an improvement as double as large compared to that when moving from 2 to 3, or from 7 to 8. For calculating bivariate correlations (number of factors that determine happiness = 1), the ordinality of SWB measures calls for application of the Spearman correlation measure, while for multivariate analyses (number of factors that determine happiness > 1) an ordered probit or logit estimator should be employed.

8.2. Ordered probit versus OLS

However, Ferrer-i-Carbonell and Frijters (2004) show for the usual 10-point life satisfaction question that the assumption of cardinality or ordinality in the econometric model does not affect 1) direction and significance of influential explanatory variables and 2) the trade-offs between the decisive determinants. In principle, however, the bias from assuming cardinality might become larger as the number of categories declines (see also Spearman correlation coefficient versus cardinality-assuming, normal correlation). Thus, while the bias for the associations with the life

satisfaction question may be negligible, it may be still substantial for the happiness question, which, traditionally, employs fewer categories by far.[35]

Neglecting the ordinal nature of the SWB measures also leads to differences in interpretation of regression coefficients, particularly when it comes to quantitative assessments. In ordered probit and logit models marginal effects have to be calculated separately, and relate to changes in probabilities of reporting a certain SWB category. The mathematical formula for such marginal effect reveals that its size is sensitive to what values the other variables in the model assume (which are held constant), e.g. whether the researcher considers a man or a woman. It is standard procedure to fix the remaining determinants at the sample mean, but any other value combination can equally be chosen (e.g. a representative person). In contrast, in OLS models that assume cardinality estimated coefficients equal marginal effects, the sizes of which are independent of what values the remaining explanatory variables take on. Using OLS also eases the calculation and interpretation of the marginal effect of two interacting variables which is important to identify heterogeneity of effects. The interpretation of these OLS marginal effects is in terms of quantitative increases in the dependent cardinal variable, the happiness measure. However, *in praxi*, policy makers should be interested in direction of influences, relative sizes and (economic) trade-offs and rather than precise quantitative assessments. Fischer (2009b) discusses the influence of institutional and macro-economic factors on subjective well-being in OECD countries, and provides examples for trade-off assessments based on relative sizes of OLS regression coefficients.[36]

8.3. Aggregation as remedy?

The problems associated with assuming cardinality when ordinality is present could be mitigated by aggregating the individual information up to a higher level. One may view the calculation of national percentages or shares of those who report certain levels of SWB as a way to deal with ordinality, as national means clearly are based on the assumption of cardinality. The resulting aggregate measures are, however, both cardinal, allowing to apply simple statistical methods of analysis.

[35] Notably, as discussed before, the same study shows that neglecting unobserved individual heterogeneity (e.g. personality traits) biases results stronger by far compared to applying OLS to an ordinal 10-scale measure.
[36] It is recommended to base relative size comparisons on standardized coefficients.

Table 5 shows for the World Values Survey 1997-2001 world sample that the cognitive component of SWB measured either as means of life satisfaction and as population shares are strongly correlated up to a level that would be termed quasi-multicollinear in a regression context. The correlation coefficients are all well above 0.9 (in absolute terms). Table 6 repeats the same exercise for 30 OECD countries, coming to qualitatively identical conclusions.

Table 5: Simple correlations between several aggregate measures of life satisfaction for 81 countries in the World Values Survey, 1997-2001

	Mean LSF	Min LSF	Max 3 LSF	Max 4 LSF
Mean LSF	1.0000			
Min LSF	-0.9170	1.0000		
Max 3 LSF	0.9591	-0.7787	1.0000	
Max 4 LSF	0.9788	-0.8475	0.9801	1.0000

Source: World Values Survey, combined 3rd and 4th waves, 1997-2001, 81 countries, 117'264 persons.
Notes: The original life satisfaction (LSF) variable ranges from 1 to 10. 'Min LSF' denotes the share of people who report any of the three lowest categories of LSF (1, 2, or 3). Max 3 (4) LSF denotes the share of people reporting any of the three (four) highest categories of life satisfaction (1, 2, or 3 (or 4)). Observations with missing life satisfaction information have been excluded.

Table 6: Simple correlations between several aggregate measures of life satisfaction for 30 OECD countries

	Mean LSF	Min LSF	Max 3 LSF	Max 4 LSF
Mean LSF	1.0000			
Min LSF	-0.9102	1.0000		
Max 3 LSF	0.9798	-0.8399	1.0000	
Max 4 LSF	0.9840	-0.9292	0.9695	1.0000

Notes: See Table 5.

9. Conclusion: What measure of subjective well-being to use ?

To wrap up, it appears that the cognitive component (*decision utility*) is (still) deemed more relevant to social policy decision-making compared to the affective component of subjective well-being (*experienced utility*) - possibly, due to path dependency, but also because of the still high costs of generating sufficient frequency in collecting affect data - despite of their advantage of a

more accurate reflection of present-time utility (losses). These high costs of collection, by the way, do not allow developing countries to participate.

In addition, optimal human behaviour in the long-run appears to relate to decisions based on the cognitive component of SWB, not the affective one - focusing on the latter causes various prediction biases. In contrast, information on decision utility may be easily obtained by exploiting already quite common household panel surveys or population census, which are repeated at least on an annual basis. A simple way to achieve fast cross-country coverage would be to incorporate questions measuring SWB in EU-SILC (European Union Statistics on Income and Living Conditions), as soon as possible.

When it comes to measuring the cognitive component of SWB (decision utility), we have also seen that the life satisfaction question-variant is to be preferred over the happiness-variant as the first is less likely to be affected by mood swings that are a source of measurement error. It also seems recommendable to apply a 10- to 11-point scale in place of a 4- to 5-point scale, because of reasons of validity, the simplicity of statistical analysis (OLS), and interpretation of results (marginal effects).

In order to allow for comparisons of national well-being over longer time spans, however, previous (national) versions of SWB questions should be continued, at least for another 30 years (one year = one data point). Instead, for analyzing variation in SWB across individuals, panels are preferred over cross-sections, and a minimum of three consecutive years is necessary for a reasonable statistical method (individual fixed effects), and 7 consecutive years for analyzing adaptation effects.

References

Allardt, E. (1976), Dimensions of Welfare in a Comparative Scandinavian Study, Acta Sociologica 19, 227-239.

Andrews, F. M., and S.B. Withey (1976). Social indicators of well-being. New York: Plenum.

Ateca-Amestoy, V. , R. Serrano-del-Rosal, and E. Vera-Toscano (2008), The leisure experience, The Journal of Socio-Economics 37, 64–78.

Bentham, J. (1789), An introduction into the principles of morals and legislation.

Bjørnskov, C. (2008), How Comparable are the Gallup World Poll Life Satisfaction Data? Forthcoming in Journal of Happiness Studies.

Blanchflower, D.G. and A.J. Oswald (2008a), Heart Rate and Happiness, presentation at the University of Zurich, 2008.

Blanchflower, D.G. and A.J. Oswald (2008b), Hypertension and happiness across nations, Journal of Health Economics 27, 218-233.

Bradburn, N.M. (1969), The structure of psychological wellbeing. Chicago: Aldine.

Brock, D.W. (1993), Life and death: Philosophical essays in biomedical ethics, Cambridge University Press, New York.

Camfield, L. and S. Skevington (2008), On Subjective Well-being and Quality of Life, Journal of Health Psychology 13, 764-775

Cantril, H. (1965). The pattern of human concerns. New Brunswick, NJ: Rutgers University Press.

Christopher, J.C. (1999), Situating Psychological Well-Being: Exploring the Cultural Roots of Its Theory and Research, Journal of Counseling & Development 77, 141-152.

Clark, A. and O. Lelkes (2009), Let us pray: religious interactions in life satisfaction, PSE working paper, Jan 2009

Clark, A., P. Frijters, P., and M. Shields (2008), Relative income, happiness and utility: an explanation for the Easterlin paradox and other puzzles, Journal of Economic Literature 46, 95-144 .

Cohen F., K.A. Kearney, L.S. Zegans, M.E. Kemeny, J.M. Neuhaus, and D.P. Stites (1999), Differential immune system changes with acute and persistent stress for optimists vs. pessimists, Brain, Behavior, and Immunity 13, 155-74.

Courard-Hauri, D. (2007), Using Monte Carlo analysis to investigate the relationship between overconsumption and uncertain access to one's personal utility function, Ecological Economics 64, 152-162.

Cummins, R.A. (1996), The domains of life satisfaction: An attempt to order chaos, Social Indicators Research 38, 303-328

Cummins, R.A. (1993), Comprehensive Quality of Life Scale for adults. Manual ComQol-A4, Deakin University, School of Psychology, Australia

Daly, M.C. and D.J. Wilson (2008), Happiness, Unhappiness, and Suicide: An Empirical Assessment. Federal Reserve Bank of San Francisco Working Paper, 2008-19.

Daly, M.C., D.J. Wilson, and N. J. Johnson (2007). Happiness Relative status and well-being: evidence from US suicide deaths, Federal Reserve Bank of San Francisco Working Paper, 2007-12.

Deaton, A. (2007), Income, Ageing, Health and Wellbeing around the World: Evidence from the Gallup World Poll, NBER Working Paper 13317.

Diener, E., R.E. Lucas, U. Schimmack, and J. Helliwell (2008), Well-Being for Public Policy, Oxford, UK: Oxford University Press.

Diener, E., R.A. Emmons, R. J. Larsen, and S. Griffin (1985), The Satisfaction With Life Scale, Journal of Personality Assessment 49, 71 – 75.

Diener, E. (1984). Subjective well-being. Psychological Bulletin 95, 542–575.

Dolan, P., T. Peasgood, and M. White (2008), Do we really know what makes us happy ? A review of the economic literature on the factors with subjective well-being, Journal of Economic Psychology 29, 94-122.

Dorn, D., J. A.V. Fischer, G. Kirchgässner, and A. Sousa-Poza (2008), Direct democracy and life satisfaction revisited: new evidence for Switzerland, Journal of Happiness Studies 9, 227-255.

Duncan, G. (2005), What Do We Mean By 'Happiness'? The Relevance of Subjective Wellbeing to Social Policy, Social Policy Journal Of New Zealand 25, 16-31

Easterlin, R. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In: Nations and Households in Economic Growth. P. A. David and M. W. Reder, Academic Press, 89-125.

Ehrhardt, J.J., W.E. Saris, and R. Veenhoven (2000), Stability of Life-satisfaction over Time, Journal of Happiness Studies 1, 177-205

Ferrer-i-Carbonell, A. (2005), Income and well-being: an empirical analysis of the comparison income effect, Journal of Public Economics 89, 997-1019.

Ferrer-i-Carbonell, A. and P. Frijters (2004), How Important is Methodology for the estimates of the determinants of Happiness?, Economic Journal 114, 641-659.

Fischer, J. A.V. (2009a), What makes people happy in OECD countries? A micro-level analysis, mimeo, Paris.

Fischer, J. A.V. (2009b), Cross-country determinants of subjective well-being in 30 OECD countries, presented at the ISQOLS 2009, July 19-23, Florence.

Fischer, J.A.V. (2009c), Happiness and age cycles – return to start…, MPRA Paper 15249, University Library of Munich, Germany, ESEM 2009 Conference Paper.

Fischer, J. A.V. and A. Sousa-Poza (2008), Personality, Job Satisfaction and Health - The Mediating Influence of Affectivity, Swiss Journal of Economics and Statistics 144, 379-435.

Frey, B.S. and A. Stutzer, 2000, Happiness, Economy and Institutions, The Economic Journal, 110 , 918-938

Fugl-Meyer, A.R., I.B. Braenholm, and K.S. Fugl-Meyer (1991), Happiness and domain-specific life satisfaction in adult northern Swedes, Clinical Rehabilitation 5, 25-33.

Hamermesh, D.S. and N.M. Soss (1974), An Economic Theory of Suicide, The Journal of Political Economy 82, 83-98

Helliwell, J.D. (2008), Life Satisfaction and Quality of Development, NBER working paper no. 14507.

Helliwell, J.D. (2007), Well-Being and Social Capital: Does Suicide Pose a Puzzle?, Social Indicators Research 81, 455-496.

Huppert, F.A. (2008), State-of-Science Review: Psychological Wellbeing: Evidence regarding its Causes & Consequences. Foresight-SR-X2 v5 stage 6, Mental Capital and Mental Wellbeing, Version 9 July 2008.

Juster, F.T., P. N. Courant, and G.D. Dow (1985), A Conceptual Framework for the Analysis of Time Allocation Data, in: F.T. Juster and F.P. Stafford (eds.) , Time, Goods and Well-Being, Ann Arbor, Michigan: Institute for Social Research, University of Michigan, 113-31.

Kahneman, D. and A.B. Krueger (2006), Developments in the Measurement of Subjective Well-Being, Journal of Economic Perspectives 20, 3-24.

Kahneman, D. and R.H. Thaler (2006), Anomalies: Utility Maximization and Experienced Utility, Journal of Economic Perspectives 20, 221-234.

Kahneman, D., B. Fredrickson, C.A. Schreiber, and D. Redelmeier (1993), When More Pain is Preferred to Less: Adding a Better End, Psychological Science 4, 401-405.

Koivumaa-Honkanen, H., R. Honkanen M. Koskenvuo, and J. Kaprio (2003), Self-reported happiness in life and suicide in ensuing 20 years, Social Psychiatry and Psychiatric Epidemiology, 38, 244-248.

Kroh, M. (2009), An Experimental Evaluation of Popular Well-Being Measures, DIW Berlin, unpublished manuscript.

Kiecolt-Glaser, J.K. , L. McGuire, T.F. Robles, and R. Glaser (2002), Psychoneuroimmunology: Psychological Influences on Immune Function and Health, Journal of Consulting and Clinical Psychology 70, 537–547

Kilpatrick F., and H. Cantril (1960), Self-anchoring scaling: A measure of individual's unique reality worlds, Journal of Individual Psychology 16, 158-170.

Layard, R. (2005), Happiness: Lessons from a New Science, London: Penguin.

Likert, R. (1932), A Technique for the Measurement of Attitudes, Archives of Psychology 140, 1–55

Lucas, R., E. Diener, and E. Suh (1996), Discriminant validity of well-being measures, Journal of Personality and Social Psychology 71, 616-28.

Lykken, D.T. and A. Tellegen (1996), Happiness is a Stochastic Phenomenon, Psychological Science 7, 186-189.

McCall, S. (1975), Quality of life, Social Indicators Research 2, 299-249

Nielsen, I. and R. Smyth (2008), Job satisfaction and response to incentives among China's urban workforce, The Journal of Socio-Economics 37, 1921–1936

Nussbaum, M.C. (2000), Women and Human Development: The Capabilities Approach, Cambridge University Press.

Mill, J.S. (1863), Utilitarianism.

Mojon-Azzi, S. and A. Sousa-Poza (2007), Hypertension And Life Satisfaction: A Comment And Replication Of Blanchflower And Oswald (2007), University of St. Gallen Department of Economics working paper series 2007-44

OECD (2007), Society at a Glance: OECD Social Indicators 2006, Paris: OECD 2007.

Oswald, A. and N. Powdthavee (2006), Does Happiness Adapt? A Longitudinal Study of Disability with Implications for Economists and Judges, Institute for the Study of Labor (IZA), IZA Discussion Papers 2208,

Ouelette-Kuntz, H. (1990), A pilot study in the use of the Quality Of Life Interview Schedule, Social Indicators Research 23, 283-298

Ouweneel, P., and R. Veenhoven (1991), Cross-National Differences in Happiness: Cultural Bias or Societal Quality?, in: Bleichrodt, N & Drenth, P.J. (eds.) Contemporary issues in cross-cultural psychology, Amsterdam: Swets & Zeitlinger 1991, 168-184

Plato, Politeia, book II

Robinson, W.S. (1950), Ecological correlations and the behavior of individuals, American Sociological Review, 30, 351-357

Rode, J.C., M.T. Rehg, J.P. Near, and J.R. Underhill (2007), The Effect of Work/Family Conflict on Intention to Quit: The Mediating Roles of Job and Life Satisfaction, Applied Research in Quality of Life 2, 65–82

Sandoe, P. (1999), Quality of life, three competing views, Ethical theory and Moral Practice 2, 11-23.

Schimmack, U., E. Diener, and S. Oishi (2002), Life-Satisfaction Is a Momentary Judgment and a Stable Personality Characteristic: The use of chronically accessible and stable sources, Journal of Personality 70, 345-384.

Schimmack, U. and S. Oishi (2005), The Influence of Chronically and Temporarily Accessible Information on Life Satisfaction Judgments, Journal of Personality and Social Psychology 89, 395–406.

Saris, W.E., T. van Wijk, and A. Scherpenzeel (1998), Validity and Reliability of Subjective Social Indicators: The effect of different measures of association 45, 173-199.

Schimmack, U., P. Krause, G.G. Wagner, and J. Schupp (2009), Stability and Change of Well Being: An Experimentally Enhanced Latent State-Trait-Error Analysis, Social Indicator Research (forthcoming), DOI: 10.1007/s11205-009-9443-8

Schimmack, U., J. Schupp, and G.G. Wagner (2008), The Influence of Environment and Personality on the Affective and Cognitive Component of Subjective Well-being, Social Indicator Research 89, 41-60

Schneider, L., and U. Schimmack (2009). Self-Informant Agreement in Well-Being Ratings: A Meta-Analysis, Social Indicator Research (forthcoming), DOI: 10.1007/s11205-009-9440-y

Schwarz, N. and F. Strack (1999), Reports of subjective well-being: Judgmental processes and their methodological implications. In: D. Kahneman, E. Diener & N. Schwarz (Eds.), Well-being: the foundations of hedonist psychology, New York: Russell Sage Foundation, 61-84.

Schwarz, N. and G.L. Clore (1983), Misattribution, and Judgements of Well-Being: Informative and Directive Functions of Affective States. Journal of Personality and Social Psychology 45, 513-523.

Sen, A. (1985), Commodities and capabilities, Amsterdam: North-Holland

Steptoe A., J. Wardle, and M. Marmot (2005), Positive affect and health-related neuroendocrine, cardiovascular, and inflammatory processes. Proceedings of the National Academy of Science of the United States 102, 6508-6512.

Stubbe, J.H., D.I. Posthuma, and E.J.C. De Geus, (2005), Heritability of life satisfaction in adults: a twin-family study, Psychological Medicine. 35, 1581-1588

Sumner, W. (1996). Welfare, happiness, and ethics. Oxford: Claredon Press.

Urry, H.L., J.B. Nitschke, I. Dolski, D.C. Jackson, K.M. Dalton, C.J. Mueller, M.A. Rosenkranz, C.D. Ryff, B.H. Singer, and R.J. Davidson (2004), Making a Life Worth Living: Neural Correlates of Well-Being, Psychological Science 15, 367–372.

van Reekum, C.M., H.L. Urry, T. Johnstone, M.E. Thurow, C.J. Frye, C.A. Jackson, H.S. Schaefer, A.L. Alexander, and R. J. Davidson (2007), Individual Differences in Amygdala and Ventromedial Prefrontal Cortex Activity are Associated with Evaluation Speed and Psychological Well-being, Journal of Cognitive Neuroscience 19, 237-248

Veenhoven, R. (2000), The four qualities of life: Ordering concepts and measures of the good life, Journal of Happiness Studies 1, 1-39.

Veenhoven, R. (2009), Happiness Scale Interval Study http://www.worlddatabaseofhappiness.eur.nl/scalestudy/scale_fp.htm (9 Jan 2009)

Weitoft, G.R., and M. Rozen (2005), Is perceived nervousness and anxiety a predictor of premature mortality and severe morbidity? A longitudinal follow up the Swedish survey of living conditions, Journal of Epidemiology and Community Health 59, 794-798.

Wanous, J.P. and M.J. Hudy (2001), Single-item reliability: a replication and extension, Organizational Research Methods 4,361-375.

Wilson, W. (1967), Correlates of avowed happiness, Psychological Bulletin 67, 294-406.

Ware, J.E. Jr. (1996), The SF-36 Health Survey, In: B. Spilker (Ed.), Quality of Life and Pharmaco-economics in Clinical Trials, Leppincott-Raven Publishers, Philadelphia USA 1996, 337-345

WHO (2001), The World Health Report 2001 - Mental Health: New Understanding, New Hope, Geneva.